Linguistic Clues to Iroquoian Prehistory

MICHAEL A. SCHILLACI, *Department of Anthropology, University of Toronto Scarborough, Toronto, Ontario, M1C 1A4, Canada. Email: schillaci@utsc.utoronto.ca*
CRAIG KOPRIS, *Waⁿdat Yaṇohšetsih, Washington, DC, USA*
SØREN WICHMANN, *Leiden University Centre of Linguistics, The Netherlands, and Laboratory of Quantitative Linguistics, Kazan Federal University, Russia*
GENEVIEVE DEWAR, *Department of Anthropology, University of Toronto Scarborough*

This paper employs a quantitative analysis of lexical data to generate a tree describing the historical relationships among Iroquoian languages. An alternative to glottochronology is used to estimate the timing of branching events within the tree. We estimate the homeland of the language family using lexical and geographic distance measures and then compare this estimate with homeland determinations in the literature. Our results suggest that Proto-Iroquoian dates to around 2624 BC, and that the Finger Lakes region of west-central New York is the most likely homeland. The results also revealed a strong relationship between linguistic dissimilarity and geographic distance, likely reflecting the isolating effects of spatial separation on the magnitude of linguistic exchange. The timing of language divergences seems to coincide with important events observable in the archaeological record, including the first evidence for the use of corn in New York and Ontario. The development of important Iroquoian cultural attributes such as the longhouse, matrilocal residence, and the intensification of agriculture all coincide with a period which saw most of the internal language divergences.

Key words: Iroquoian, archaeology, language, linguistics, ASJP

Linguistic data have the potential to contribute in a unique way to our understanding of prehistory, both regionally and globally. In North America, there are numerous examples in the literature of linguistic data contributing to archaeological reconstructions of culture history (e.g., Davis 1959; Fiedel 1987, 1990, 1991; Ortman 2012). Most archaeological reconstructions of North American prehistory using linguistic data have relied on glottochronological estimates of language divergence based on the proportion of shared cognates, which are determined through analysis of sound correspondences, phonetic similarity, and semantic affinity. We attempt here to contribute to the study of Iroquoian prehistory by estimating historical relationships among Iroquoian languages and the timing of their divergence through quantitative analyses of lexical data. In addition, we identify the regional homeland of the Iroquoian language family using geographic and lexical data. Although we explore possible temporal correspon-

Submitted April 24, 2016; accepted October 25, 2016; published online July 21, 2017.
Journal of Anthropological Research (Fall 2017). 0091-7710/2017/7303-0004$10.00

dences between the timing of language divergence and important developments in Iroquoian prehistory as reflected in the archaeological record, our objective is not to evaluate directly specific models regarding Iroquoian origins presented by archaeologists, nor is it our aim to tie specific archaeological cultures or traditions to Iroquoian protolanguages. Instead, our intention is to present a chronological and geographic context based on linguistic data that may inform archaeological inquiry.

The origin of Iroquoian-speaking peoples has received considerable attention in the academic literature, particularly by archaeologists interested in linking archaeological cultures with Iroquoian ethnolinguistic identity. Most of the archaeological literature has focused on Northern Iroquoian origins and prehistory, rather than the entire language family. Until recently, two prevailing hypotheses regarding Northern Iroquoian origins have been put forward by archaeologists: (1) the migration hypothesis and (2) the in situ hypothesis. The specific details and historical development of these hypotheses have been reviewed thoroughly elsewhere (e.g., Birch 2015, Crawford and Smith 1996; Hart 2001; Martin 2008; Snow 1995; Trigger 1970; Warrick 2000; Williamson 2014). Briefly, migration hypotheses place the origin of Northern-Iroquoian-speaking peoples outside of New York and Ontario. Stothers (1977) placed their origins in the Midwestern United States, perhaps within the Hopewell cultures of the Middle Woodland period (300 BC–AD 500),[1] whereas Snow (1995) suggested Northern Iroquoian origins could be found within the Clemson's Island culture (AD 750–1300) of central Pennsylvania (also see Snow 1996). In both of these scenarios, the migration from the south or southwest brought Northern-Iroquoian-speaking peoples to what is now New York, Ontario, and Quebec by the time of European contact. The in situ hypothesis, on the other hand, posits an autochthonous development of Northern-Iroquoian-speaking populations in the southern Great Lakes region, including much of southern Ontario and western and central New York, as well as the shores of the St. Lawrence River near its western end. In New York, the timing of this development may have coincided with the emergence of the Late Woodland period (ca. AD 500–1600), though deeper temporal roots in the Middle Woodland (ca. 300 BC–AD 500), or even Late Archaic (4000–1000 BC) periods has been suggested (e.g., Wright 1984). In Ontario, this in situ development coincides with the earliest appearance of maize (Crawford and Smith 1996) associated with the Princess Point complex (AD 500–1000) of the initial Late Woodland period of southern Ontario. Although not well described in the literature for either the in situ or migration hypotheses, the origins of Southern Iroquoian (Cherokee) and the Tuscarora-Nottoway-Meherrin linguistic groupings would be attributable either to separate migrations to the southeastern United States at some point in prehistory, assuming the in situ model, or might have been surviving representatives of an earlier, larger distribution of Proto-Iroquoian and/or Proto–Northern Iroquoian speakers, with a possibility of multiple migratory events within the same general geographic domain. This latter scenario presents the possibility of an Iroquoian homeland to the south of present-day Northern-Iroquoian-speaking groups, perhaps within the central Appalachian region.

Although the debate regarding the origins of Iroquoian-speaking peoples as a collective ethnolinguistic entity continues, the current discourse on Iroquoian prehistory recognizes that the historical process of Iroquoian ethnogenesis was likely complex and heterogeneous (Birch 2015:270). The use of ethnogenesis as a concept in archaeology recognizes that the development of a collective ethnic identity often comprises multiple intersecting components including shared biology, culture, language, and history. Data sets relating to material culture, language, and genetic relationships, however, may not overlap or provide similar temporal and geographic patterning (see Chrisomalis and Trigger 2004). This reflects the fact that biological populations and their culture and languages do not always develop, or evolve, as bounded ethnic packages (Ortman 2012:2–3; also see Bateman et al. 1990; Campbell 2015). Similarly, shared cultures and languages do not necessarily develop or evolve in the same way among populations through time. For example, Snow (2009:9) has pointed out that the overall continuity of a language over time is often better than linguistic continuity of a given speech community, assuming multiple speech communities comprise that language. It is thus important not to confuse a language with the community that speaks it. As was recently asserted by Ortman (2012), the key to solving this puzzle may be to view all of these data sets separately and employ methods that allow culture, language, genes, and people to derive from different sources, and change in different ways. Ethnogenesis, therefore, may sometimes involve multiple social and biological processes involving multiple peoples and cultures from different geographic regions (Ortman 2012; Schillaci and Bustard 2010). The process of ethnogenesis eventually leads to the development of cohesive and integrative cultural mechanisms, including rituals and various forms of social organization, as well as a collective identity. That process is likely dynamic and heterogeneous, involving both regional (in situ) and extraregional (migration) components. Also, as pointed out by Birch (2015:306), the process of Iroquoian ethnogenesis was not restricted to a specific event or point in time marking Iroquoian origins, but rather continued throughout the precontact and contact periods, with cultural change accelerating during periods of coalescence, migration, population expansion, and incorporation. Given the likely complex and heterogeneous nature of Iroquoian ethnogenesis, it is not surprising that recent attempts to identify Iroquoian ethnic groups archaeologically using ceramic design variation have been unsuccessful (Hart and Engelbrecht 2012).

THE IROQUOIAN LANGUAGE FAMILY

Nine languages currently comprise the Iroquoian language family of eastern North America, with an additional eight or more languages spoken at the time of European contact. These languages are divided into two main groupings: Southern Iroquoian, with Cherokee as its sole representative, and Northern Iroquoian, represented by Seneca, Cayuga, Onondaga, Oneida, Mohawk, Wendat (Huron), Wyandot (Petun), and Tuscarora, as well as historical languages no longer spoken, such as Susquehannock, Nottoway, Meherrin, Neutral, Wenro, Erie, and the Laurentian languages (Hoche-

lagan, Stadaconan). Northern Iroquoian divides into two subgroups. The first, which has been termed Coast Iroquoian by Kopris (2001), is made up of Tuscarora and two historically documented languages, Nottoway and Meherrin (Mithun 1984). The second, and largest, of these subgroupings, which has been referred to by Mithun (1984) as Lake Iroquoian, includes the Huronian languages currently represented by Wendat and Wyandot, and the languages of the Five Nations of New York represented by Seneca, Cayuga, Onondaga, Oneida, and Mohawk. As pointed out by Mithun (1984), because the documentation of historical languages no longer spoken is fragmentary and incomplete, their position within the language group is uncertain. Nonetheless, the available evidence seems to indicate that Nottoway and Meherrin likely group with Tuscarora (Rudes 1981), while Susquehannock, Neutral, Wenro, Erie, and Laurentian fall within the Lake Iroquoian grouping (Mithun 1981, 1984).

The structure of historical relationships among Iroquoian languages has been examined in a variety of ways, including counts of shared cognates (Blin-Lagarde 1972; Hoffman 1959; Julian 2010; Lounsbury 1961), measurements of mutual intelligibility (Hickerson et al. 1952), and the identification of phonological and morphological innovations (Chafe and Foster 1981). With the exception of the analysis of phonological innovations by Chafe and Foster (1981), which differs only in its placement of Cayuga, all of these methods have produced a structure of historical relationships that reflect the widely accepted groupings described in the previous paragraph (cf. Lounsbury 1978: fig. 1; Mithun 1984: fig. 15.2). Language trees generated using the percent shared cognates presented by Hoffman (1959) (based on 48 words), Blin-Lagarde (1972) (111 words from the Swadesh 200-word list), and Julian (2010) (Swadesh 100-word list) are largely similar (Figure 1). Accordingly, the numerical values for percent cognates shared for those languages common to all three studies (Table 1) are highly correlated, ranging from $r_s=0.893$ to $r_s=0.952$ (where r_s is Spearman's correlation coefficient; for all three correlations, $p<0.001$).

The language tree based on values presented by Hoffman (1959) differs by including shared cognates for several historically documented languages (Laurentian, Susquehannock, and Nottoway) not included by Julian (2010) in his analysis. The tree based on data from Hoffman (1959) also differs by grouping Mohawk with Onondaga, rather than Oneida, a finding that seems to be inconsistent with most descriptions of historical relationships presented in the literature. The language tree generated from the pair-wise estimates of percent mutual intelligibility among Five Nations Iroquoian languages presented by Hickerson et al. (1952) is identical to the tree based on shared cognate percentages presented in Julian (2010). The estimates of mutual intelligibility are also correlated with the cognate percentage values presented by Julian (2010) ($r_s=0.810$, $p<0.001$), Hoffman (1959) ($r_s=0.789$, $p<0.001$), and Blin-Lagarde (1972) ($r_s=0.879$, $p<0.001$), suggesting that percent shared cognates is a relatively good predictor of how well members of a speech community understand a closely related language or dialect. A study of Sinitic varieties (Tang and van Heuven 2009:724) found similar correlations when word intelligibility and sentence intelligi-

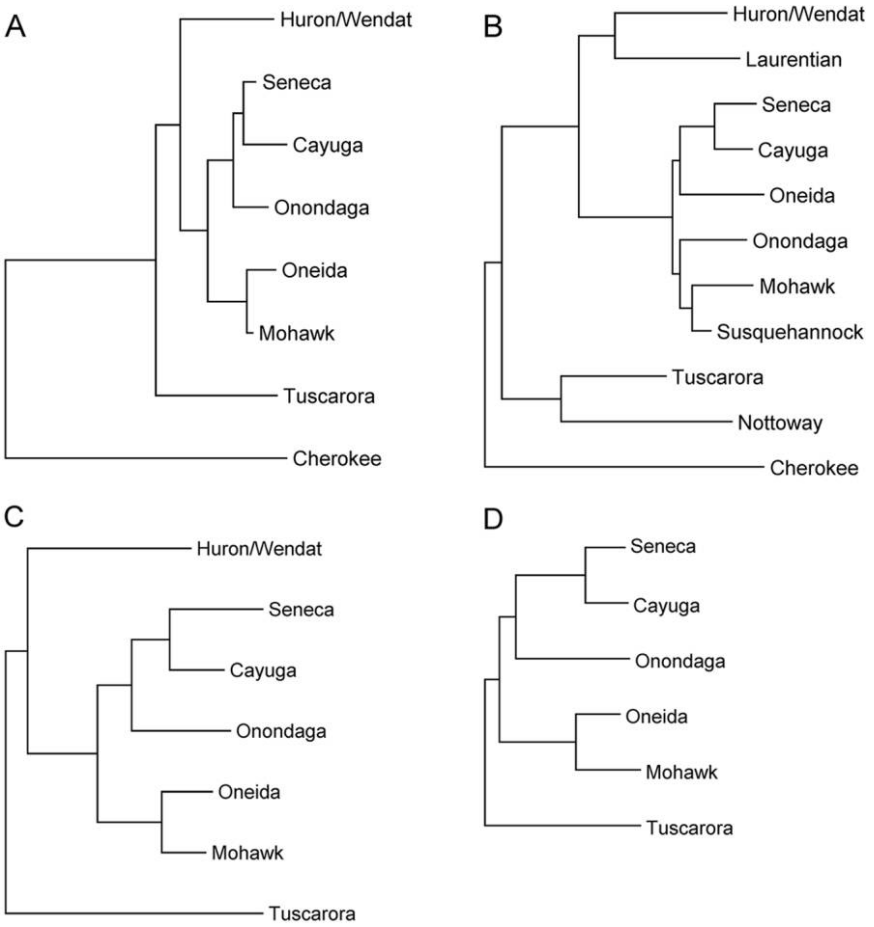

Figure 1. Neighbor-joining language tree based on (A) percent shared cognates presented in Julian (2010), (B) percent shared cognates presented in Hoffman (1959), (C) percent shared cognates presented in Blin-Lagarde (1972), and (D) percent mutual intelligibility presented in Hickerson et al. (1952). Percent shared cognate and percent mutual intelligibility values were converted to measures of dissimilarity before neighbor-joining analysis. Languages within the tree are arranged geographically, with the northernmost languages at the top of the tree and the southernmost languages at the bottom. Languages comprising subclades are arranged west (top) to east (bottom).

bility were correlated with lexical similarity as measured by cognate sharing (r=0.788 and r=0.746, respectively). But when it comes to very close dialects, another study shows that phonetic distance is a better predictor of intelligibility than lexical distance (Gooskens et al. 2008).

The dating of protolanguages, or the timing of the divergence of various Iroquoian languages and language groupings, has been attempted using glottochronology.

Table 1. Matrices describing the percent shared cognates, and percent mutual intelligibility. In matrix A, shared cognate values presented in Julian (2010) are listed above the diagonal, while those presented in Hoffman (1959) are listed below the diagonal. In matrix B, shared cognate values presented in Blin-Lagarde (1972) are listed below the diagonal, while those presented by Lounsbury (1961) are listed above the diagonal. Percent mutual intelligibility values presented in Hickerson et al. (1952) are listed in matrix C.

A. Shared cognates (Julian 2010; Hoffman 1959)

	Cher	Tusc	Moh	One	Ono	Cay	Sen	Hur	Laur	Sus	Nott
Cherokee (Cher)	—	20	23	22	24	19	21	19			
Tuscarora (Tusc)	40	—	69	65	66	65	65	67			
Mohawk (Moh)	30	45	—	95	85	83	86	76			
Oneida (One)	30	50	80	—	80	82	81	72			
Onondaga (Ono)	30	45	80	80	—	86	93	74			
Cayuga (Cay)	30	45	85	80	80	—	91	71			
Seneca (Sen)	35	55	80	80	80	90	—	76			
Huron/Wendat (Hur)	40	45	60	55	60	60	60	—			
Laurentian (Laur)	35	45	50	55	60	45	60	70	—		
Susquehannock (Sus)	25	60	90	80	90	85	80	65	80	—	
Nottoway (Nott)	35	65	35	40	40	35	45	35	40	50	—

B. Shared cognates (Blin-Lagarde 1972; Lounsbury 1961)

	Tusc	Moh	One	Ono	Cay	Sen	Hur
Tuscarora	—		59.4			50	
Mohawk	59	—					
Oneida	55	91	—		77.8	64.8	
Onondaga	56	77	77	—			
Cayuga	54	78	79	80	—	71.6	
Seneca	49	72	72	80	86	—	
Huron/Wendat	58	68	65	63	68	64	—

C. Mutual intelligibility

	Tusc	Moh	One	Ono	Cay	Sen
Tuscarora	—					
Mohawk	0.00	—				
Oneida	0.00	64.0	—			
Onondaga	0.00	8.25	19.0	—		
Cayuga	2.50	15.25	15.75	21.75	—	
Seneca	0.62	4.75	20.75	29.25	72.75	—

Glottochronology (Lees 1953; Swadesh 1950, 1955) is a quantitative method for estimating when two languages split or diverged based on the assumption that the vocabulary of a language changes over time at an approximately constant rate, which is the same for all languages (Davis 1959). The method first estimates the number of shared cognates for all possible pairs of languages from a standard list of basic vocabulary items developed by Swadesh (1950, 1955). For each pair of languages, the log of the percentage of shared cognates is then divided by the log of the squared retention rate to yield an estimate in millennia for the split of the two languages. Glottochronology has been widely criticized for its use of a constant and universal rate of language change. Because cognate determinations can vary among linguists, glottochronological estimates for language split dates sometimes vary from study to study.

The first glottochronological estimates for the Iroquoian language family were provided by Lounsbury (1961) based on cognates determined from the Swadesh 200-word list for Cherokee, Tuscarora, Seneca, Cayuga, and Oneida. In that study, Lounsbury found that Cherokee (Southern Iroquoian) shared between 37.8% and 34.3% of cognates with Northern Iroquoian languages, indicating that these two groupings split between 3,500 and 3,800 years ago. He estimates that the split between Tuscarora and the Five Nations languages occurred approximately 1,900 to 2,400 years ago, and that the split among the Five Nations languages occurred 1,200 to 1,500 years ago. Blin-Lagarde (1972), using 111 words from the Swadesh 200-word list, determined the percentage of shared cognates for groupings of languages in an iterative fashion, with each iteration omitting one of the seven Northern Iroquoian languages included in the study. She then provided glottochronological estimates for the *tronc commun* or common ancestor of these groupings which corresponded to the divergence date for the language omitted. Using this method Blin-Lagarde estimated that Tuscarora split from the remaining Northern Iroquoian languages sometime between 1,800 and 2,400 years ago (~400 BC–AD 200), depending on the retention rate (r) that was used (i.e., 0.805 or 0.85). She also estimated that Wendat (Huron) diverged from Five Nations Iroquois between 1,000 and 1,300 years ago (~AD 700–1000), and that the initial divergences among Five Nations language groupings occurred between about 850 and 1100 years ago (AD 900–1150). More recently, using a modified application of glottochronology, Brown (2010) estimated Southern and Northern Iroquoian to have split 2,700 years ago. Brown (2006, 2010) also estimated that Tuscarora split from the other Northern Iroquoian languages 2,200 years ago, the Five Nations languages began to split 1,000 years ago, and the two Huronian languages (Wendat and Wyandot) split from each other around 740 years ago.

Most recently, Holman et al. (2011) provided divergence dates for Iroquoian language groupings using a computerized alternative to glottochronology called Automated Similarity Judgment Program (ASJP). The ASJP generates dissimilarity measures among languages or dialects derived from lexical data (words), rather than the proportion of shared cognates. The ASJP dates for the Iroquoian language family are consistently older than glottochronological dates. For example ASJP estimates the

split between Southern and Northern Iroquoian to have occurred 4,855 years ago, which is between 2,155 and 1,055 years earlier than the glottochronological estimates by Lounsbury (1961) and Brown (2010). Similarly, the ASJP date for the split between Tuscarora and the Five Nations languages (3,176 years) is between 1,276 and 776 years earlier than the glottochronological estimates.

At the time of European contact, the sole Southern Iroquoian language, Cherokee, was presumably spoken within the southeastern United States, primarily in what are now South Carolina, North Carolina, Georgia, and Tennessee. Northern Iroquoian languages, including those that are extinct, were spoken in New York (Five Nations, Wenro, Erie, Laurentian), Pennsylvania (Susquehannock, Erie), southern Ontario (Neutral, Wendat, Wyandot), western Quebec along the shores of the St. Lawrence River (Laurentian), North Carolina (Tuscarora), and southern Virginia (Nottoway, Meherrin) (see Boyce 1978:282; Trigger 1978). There has been considerable movement of Iroquoian language speakers since European contact, owing primarily to warfare and forced relocations as well as voluntary migration. Consequently, Iroquoian-speaking groups currently reside outside their historical ranges summarized above. For example, there are Five Nations groups residing in Ontario and Quebec, and even Wisconsin. Similarly, there are Cherokee-speaking groups in Oklahoma, and Huronian-speaking groups in Quebec (Wendat), Michigan, Kansas, and Oklahoma (Wyandot). Although the geographic ranges of Iroquoian groups are reasonably well known for the period immediately following European contact, their precontact ranges are not, and have been the subject of considerable interest for anthropologists, particularly archaeologists (see Williamson 2014).

METHODS

Words corresponding to a 40-word subset of the Swadesh 100-word list (Swadesh 1955) (Table 2) for nine languages of the Iroquoian language family were gathered from the literature and from experts. This list of 40 words has been found to yield lexicostatistical results at least as accurate as the full 100-word list as determined by their correlation with language classifications by specialists (Holman et al. 2008). Using the shorter list allowed us to obtain complete data sets for all well-documented Iroquoian languages. The primary sources for the word lists are as follows: Cherokee (Feeling 1975), Cayuga (Froman et al. 2002), Mohawk (Michelson 1973), Oneida (Michelson and Doxtator 2002), Onondaga (Woodbury 2003), Seneca (Chafe 1967), Tuscarora (Rudes 1999), Wendat (Yawenda Project),[2] and Wyandot (Kopris 2001, unpublished data). After differences among published sources in orthography were reconciled,[3] the words were transcribed into a standard orthography called ASJPcode (Holman et al. 2011) and deposited in the ASJP database (Wichmann et al. 2016).

A distance matrix of pair-wise measures of lexical dissimilarity (LDND) among the Iroquoian languages was generated from the 40-word subset of the Swadesh 100-word list using ASJP (Holman et al. 2011). This measure is based on a Levenshtein distance (LD), which is defined as the minimum number of successive changes needed to

Table 2. 40-word subset of the Swadesh 100-word list (Swadesh 1955) used by the ASJP to generate the LDND measures of lexical dissimilarity among languages.

Swadesh No.	Word	Swadesh No.	Word
1	I	47	knee
2	you	48	hand
3	we	51	breast
11	one	53	liver
12	two	54	drink
18	person	57	see
19	fish	58	hear
21	dog	61	die
22	louse	66	come
23	tree	72	sun
25	leaf	74	star
28	skin	75	water
30	blood	77	stone
31	bone	82	fire
34	horn	85	path
39	ear	86	mountain
40	eye	92	night
41	nose	95	full
43	tooth	96	new
44	tongue	100	name

change one word to another, where each change is a deletion, insertion, or substitution of a symbol representing a class of speech sound (Holman et al. 2011:843). The resulting value is then normalized by dividing the LD by the number of symbols of the longer of the two words. This results in a normalized Levenshtein distance (LDN) that corrects for differences in word length. An LDN divided (LDND) is then calculated by dividing the average LDN for all the word pairs involving the same meaning by the average LDN for all the word pairs involving different meanings (Holman et al. 2011:843).

Using the distance matrix we generated neighbor-joining trees (Saitou and Nei 1987) using MEGA 6.05 (Tamura et al. 2013). These trees can be interpreted as representing phylogenetic, or historical, relationships among languages. We estimated the date of each internal node within the neighbor-joining tree based on the lexical distances. These nodes correspond to the point at which a parent protolanguage splits, or diverges, into two daughter languages. Divergence dates were estimated using a computerized alternative to glottochronology developed by the ASJP consortium (see Holman et al. 2011). Briefly, this automated method estimates the time since diver-

gence t based on lexical similarity s using the formula $t = (\log s - \log s_0)/2\log r$, where $s=1-$LDND and s_0 is the average lexical similarity within time-zero languages (i.e., when $t=0$), and r is the average proportion of lexical similarity retained after a standard period of time (Holman et al. 2011: 844). Both s_0 and r are constants that must be estimated using linear regression of known dates of the breakup of different language groups and the average measured similarities of languages within the group. Holman et al. use 52 calibration points from language groups from around the world whose diversification dates are known from history, epigraphy, or archaeology. The constant s_0 corresponds to the point on the similarity axis where the regression line hits at time 0.

The ASJP methodology based on lexical similarity has several potential advantages over standard glottochronological methods based on cognates. First, glottochronology relied on only a few calibration points, mainly from languages within a single family (Indo-European), whereas ASJP chronology is calibrated using 52 points taken from languages all over the world, allowing its developers to verify the assumption that the rate of lexical change is at least sufficiently constant to serve as a chronometer. Second, the ASJP methodology does not rely on analyst determinations of similarity. Third, whereas traditional glottochronology (Lees 1953) assumes that languages, at their point of splitting up, have no internal differences among speakers, the method of Holman et al. takes into account the well-known fact that no two speakers have identical lects. An evaluation of this method by its developers indicated a margin of uncertainty of approximately ±29% (Holman et al. 2011). This is a highly conservative estimate since only some of the uncertainty derives from variation in the rate of lexical change: a part of the uncertainty derives from dates given by archaeologists, historians, and epigraphers to the breakup times of the 52 language groups used by Holman et al. to calibrate their dating method. Thus, in reality the margin of error would be less than ±29%.

In addition to dating language splits within the Iroquoian language family, we estimated the geographic homeland of Proto-Iroquoian and Proto–Northern Iroquoian using a computer-automated tool which is based on an index computed from lexical dissimilarities and geographic distances among languages (Wichmann et al. 2010). This method for identifying homelands represents a systematic implementation of Sapir's (1916) idea that the homeland of the protolanguage of a family is the location of initial divergence, and that the earliest divergence will subsequently lead to maximal diversity in and around the original homeland (Wichmann et al. 2010).

The following summarizes the methodology detailed in Wichmann et al. (2010: 248). An initial assumption is that the homeland corresponds to the location of one of the current languages. Next, the language with the highest *diversity index* is identified and its location identified as the homeland. The diversity index of a language is the average of the lexical distances between this language and all other languages in the family divided by the average of the geographical distances between the language and all the other languages in the family. A language with close proximity to lan-

guages that are quite distant lexically will have a high index, and a language whose geographical neighbors are not particularly close and not particularly different will have a low index. As pointed out by Wichmann et al. (2010:248), this method works best when movement into new areas resembles a random walk and may be upset by directed, long-distance migrations of speakers of the protolanguage.

For this method we use the LDND as our measure of lexical distance, and straight-line distances (km) among the geographic centers of Iroquoian languages measured using the ruler tool with mouse navigation in Google™ Earth (Figure 2). A geographic center represents the approximate areal midpoint of the historical occupation range of a given Iroquoian speech community as determined visually from published maps (see Boyce 1978:282; Trigger 1978:xi). In order to navigate around Lake Ontario, distances to the Huronian languages were measured using three waypoints positioned at (1) the easternmost extension of Lake Erie, (2) the westernmost extension of Lake Ontario, and (3) on Wolfe Island, near the northeastern end of Lake Ontario.

We used the measured geographic distances among language centers to test a generalized isolation-by-distance model borrowed from the field of population genetics (Wright 1943). Here, this model predicts that linguistic divergence among populations will be proportional to geographic distances owing to the isolating effects of spatial separation on the magnitude of linguistic exchange. So, in other terms, languages without linguistic exchange will drift apart, the magnitude of which varies as a function of spatial separation (i.e., geographic distance). However, because temporal separation without linguistic exchange can have the same effect on language divergence, a true test of this model requires an estimate of the correlation between temporal and lexical distances (Konigsberg 1990). Because we do not have an independent estimate of temporal separation among languages, we assume that temporal separation is proportional to geographic distances. For our study, the relationship between lexical and geographic distance matrices was examined using simple least-squares regression and a nonparametric Mantel test (Mantel 1967) using MANTEL 3.1.[4] As discussed in Holman et al. (2015), a significant relationship between geographic and lexical distances for unrelated languages must be due to diffusion. For closely related languages, such as those used in our study, a relationship between geographic and lexical distances would be due primarily to people dispersing and becoming isolated and linguistically different over time. Nonetheless, diffusion would likely contribute to the relationship between geographic and linguistic distances among related languages to some extent, with the amount of diffusion being hard to estimate.

RESULTS

The lexical (LDND) distances presented in Table 3 were significantly correlated with measures of cognate dissimilarity based on the data presented in Hoffman (1959) (Mantel test, $r_s=0.942$, $p=0.045$), Julian (2010) (Mantel test, $r_s=0.923$, $p=0.043$), and Blin-Lagarde (1972) (Mantel test, $r_s=0.947$, $p=0.002$). We averaged the Wendat and Wyandot LDND distances to represent "Huron" in the comparisons with cog-

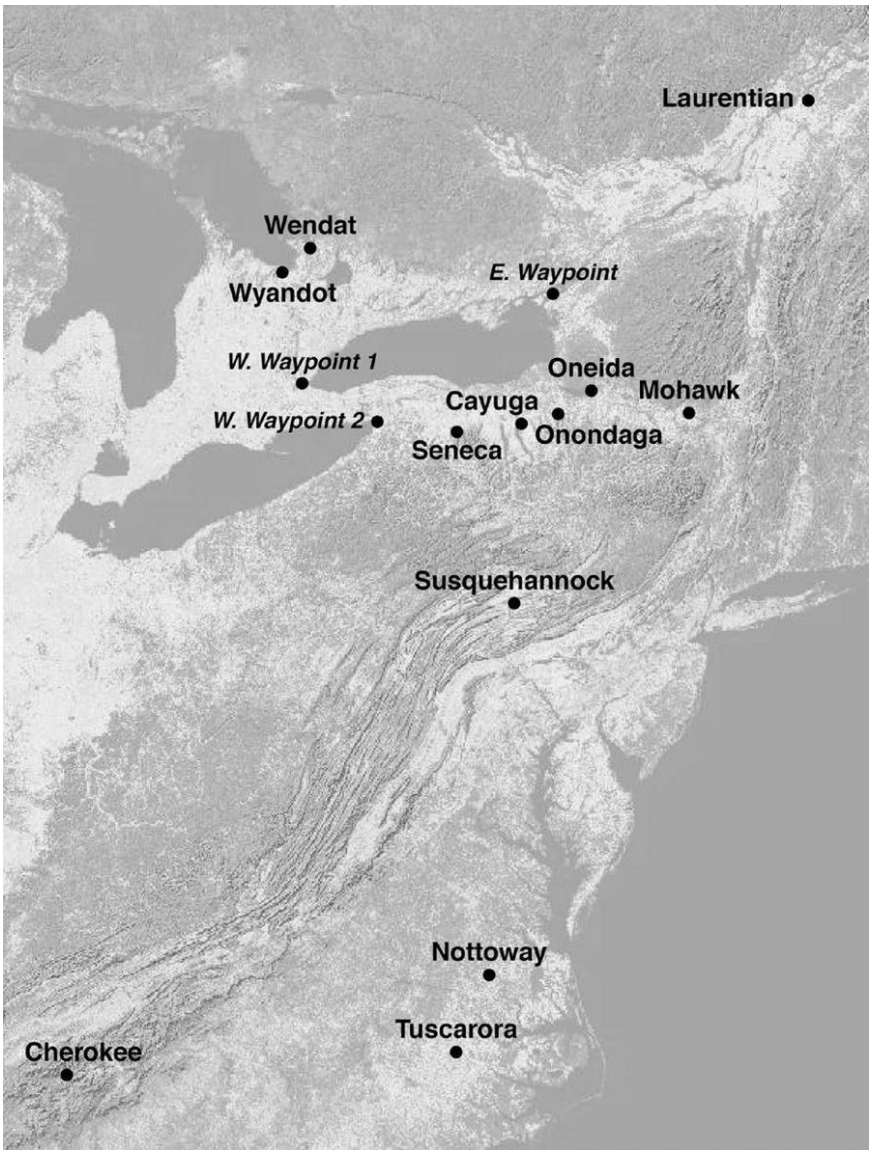

Figure 2. Map of the approximate geographic centers of the Iroquoian languages.

nate data. The LDND distances were also significantly correlated (Mantel test, r_s=0.869, p=0.013) with the intelligibility data presented in Hickerson et al. (1952) after being converted to a dissimilarity measure. The tree generated from the lexical distances (Figure 3) corresponds closely with the structure of Iroquoian language relationships estimated from the percentage of shared cognates and the pattern of language intelligibility presented earlier. As expected, our results place the sole Southern Iroquoian

Table 3. Lexical distances (LDND) among Iroquoian languages are listed below the shaded diagonal. Mean LDND distances among languages are listed for each language in the shaded diagonal. Geographic distances (km) are listed above the diagonal. Last column lists the mean geographic distance for each language.

	Cay	Moh	One	Ono	Sen	Tus	Wen	Wya	Che	Mean
Cayuga (Cay)	65.02	200	91	45	78	852	423	397	998	385.5
Mohawk (Moh)	55.98	63.95	119	156	277	852	499	529	1140	471.5
Oneida (One)	62.23	39.20	66.18	48	167	844	409	439	1080	399.6
Onondaga (Ono)	47.55	50.97	49.55	62.58	123	809	428	458	1033	387.5
Seneca (Sen)	45.43	61.26	62.86	45.20	66.76	778	351	325	945	380.5
Tuscarora (Tus)	76.90	71.21	77.07	73.29	76.46	79.36	1234	1264	513	893.3
Wendat (Wen)	69.17	69.72	72.17	71.34	76.55	82.38	72.06	43	1167	569.3
Wyandot (Wya)	67.34	67.97	72.62	67.36	70.45	80.63	39.87	70.37	1141	574.5
Cherokee (Che)	95.57	95.31	93.74	95.38	95.84	96.88	95.26	96.71	95.59	1002.1

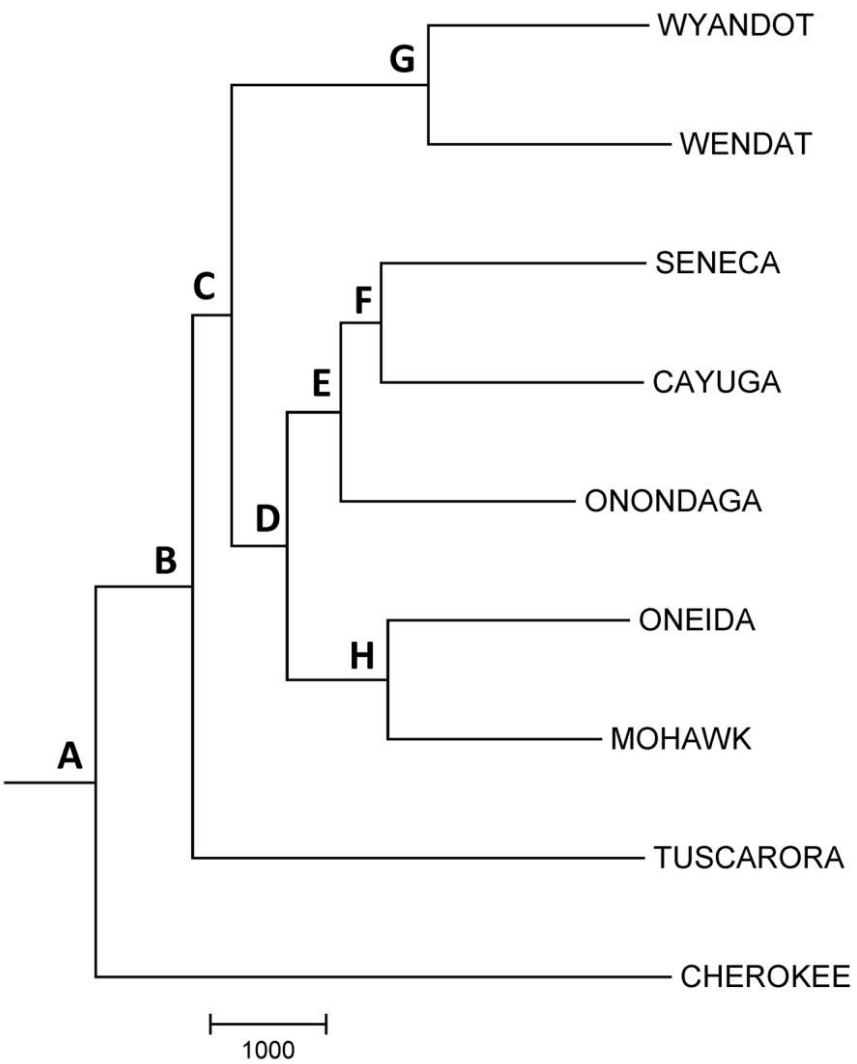

Figure 3. Neighbor-joining language tree describing the historical relationships among Iroquoian languages based on the LDND lexical distances. Nodes (branching points) are labeled A–H. Dates for labeled nodes are presented in Table 4. Languages within the tree are arranged geographically with the northernmost languages at the top of the tree and the southernmost languages at the bottom. Languages comprising subclades are arranged west (top) to east (bottom).

language, Cherokee, as a sister to the large Northern Iroquoian grouping, or clade. Within this larger clade are a series of sequential language splits. First is the divergence of Tuscarora from a clade composed of the Five Nations languages and the Huronian languages of Wendat and Wyandot (Lake Iroquoian). This initial split is followed by

the divergence of the Huronian languages from a clade composed of the Five Nations languages. Two subclades comprise this Five Nations clade, one containing Cayuga, Seneca, and Onondaga, and the second containing Mohawk and Oneida. While our results are also in general agreement with expert opinion (e.g., Lounsbury 1978: fig. 1; Mithun 1984: fig. 15.2), it is important to note that they differ from the structure of Iroquoian languages derived from an analysis of phonological and morphological innovations (Chafe and Foster 1981) which groups Cayuga with Tuscarora.

Our estimated language divergence dates based on the ASJP methodology, and the estimated dates when protolanguages were spoken, are presented in Table 4. Protolanguages span the temporal increment between clade or language divergence dates. For example, the divergence of Proto-Southern and -Northern Iroquoian, initiating the differentiation of the Iroquoian language family, dates to 4,624 ± 1,341 years ago (ya) (3965–1283 BC, after including the ±29% uncertainty interval and setting the present to the year 2000). Proto-Iroquoian would have been spoken for some unknown period of time before the divergence of Proto-Southern and -Northern Iroquoian. The Northern Iroquoian language clade, or grouping, is defined by the divergence of Proto-Tuscarora from Proto–Lake Iroquoian made up of the remaining Northern Iroquoian languages, which occurred approximately 2,100 ± 609 ya (709 BC–AD 509). This means that Proto–Northern Iroquoian was spoken between 4,624 ± 1,341 and 2,100 ± 609 ya (i.e., 3965 BC and AD 509). Given this large time span, it is likely this protolanguage developed into several lineages, only one of which is preserved today. Within the Lake Iroquoian clade, the Proto-Huronian diverged from the Proto–Five Nations clade approximately 1,730 ± 502 ya (232 BC–AD 772), while the divergence of the Five Nations languages began 1,163 ± 337 ya (AD 500–1174). The individual Five Nations languages all diverged over an approximately 200-year period between 822 ± 238 ya (Proto-Onondaga split from Proto-Seneca/Cayuga, AD 940–1416) and 630 ± 183 ya (Proto-Mohawk–Proto-Oneida split, AD 1187–1553). Within the Huronian language clade, Wendat and Wyandot diverged from each other about 647 ± 188 ya (AD 1165–1541), at about the same time the Five Nations languages were diverging from each other.

Our estimates for language divergence times are compared with glottochronological estimates derived from percent shared cognate values in Table 5. Our estimates are also supported by the pattern of shared cognates for 'bow' and 'smoking pipe' (Table 6), and for 'corn' (Tables 7 and 8) and 'bean' (Table 9). The emergence of the bow in the study area has been dated to approximately AD 500–700 in the Late Woodland period (Blitz 1988; but see Snarey and Ellis 2008), whereas smoking pipes are commonly documented on Early Woodland sites (1000 BC–AD 200) and were likely in use even earlier, during the Late Archaic (ca. 4000–1000 BC) (Rafferty 2016). All Northern Iroquoian languages except Tuscarora share a cognate for 'bow', which suggests that Tuscarora diverged before the appearance of the bow, a finding that is consistent with our estimated divergence for Tuscarora sometime during 709 BC–AD 509. As might be expected based on our estimate for the divergence of Southern and

Table 4. Estimated divergence dates for Iroquoian proto-languages, and corresponding dates for proto-languages. Error estimates (±) correspond to a margin of uncertainty of approximately ± 29% (Holman et al. 2011)

(Proto) Language split	Node ID†	Years Ago	Date	Archaeological Period
Southern–Northern Iroquoian	A	4,624 ± 1,341	3965–1283 BC	Late Archaic
Tuscarora–(Wendat, Wyandot, Five Nations)	B	2,100 ± 609	709 BC–AD 509	Middle Woodland
(Wendat, Wyandot)–Five Nations	C	1,730 ± 502	232 BC–AD 772	Mid-Late Woodland
(Seneca, Cayuga, Onondaga)–(Mohawk, Oneida)	D	1,163 ± 337	AD 500–1174	Late Woodland
Onondaga–(Seneca, Cayuga)	E	822 ± 238	AD 940–1416	Late Woodland
Seneca–Cayuga	F	795 ± 231	AD 974–1436	Late Woodland
Wendat–Wyandot	G	647 ± 188	AD 1165–1541	Late Woodland
Mohawk–Oneida	H	630 ± 183	AD 1187–1553	Late Woodland
Proto-Iroquoian	?→A	?–4,624	?–2624 BC‡	Late Archaic
Proto–Northern Iroquoian	A↔B	4,624–2,100	2624–100 BC‡	Early Woodland
Proto–Lake Iroquoian	B↔C	2,100–1,730	100 BC–AD 270‡	Middle Woodland
Proto–Five Nations	C↔D	1,730–1,163	AD 270–837‡	Mid-Late Woodland
Proto-Huronian	C↔G	1,730–647	AD 270–1353‡	Mid-Late Woodland
Proto-Oneida/Mohawk	D↔H	1,163–630	AD 837–1370‡	Late Woodland
Proto-Onondaga/Cayuga/Seneca	D↔E	1,163–822	AD 837–1178‡	Late Woodland
Proto-Seneca/Cayuga	E↔F	822–795	AD 1178–1205‡	Late Woodland

Note: The present study makes use of a more complete data set for the Iroquoian language family than that used by Holman et al. (2011), both in terms of the number of words available and the languages included in the study.

†Corresponds to lettered nodes on Figure 3.

‡Based on point estimates for nodes and does not include the ± 29% uncertainty interval.

Table 5. Comparison of time to divergence estimates (years ago) derived from percent shared cognates and lexical data.

Language split	Present study	Lounsbury 1961	Blin-Lagarde 1972	Brown 2010	Holman et al. 2011
Southern–Northern Iroquoian	5,965–3,283	3,800–3,500		2,700	4,855
Tuscarora–Lake Iroquoian	2,709–1,491	2,400–1,900	2,400–1,800	2,200	3,176
Huronian–Five Nations	2,232–1,228		1,300–1,000		
Five Nations	1,500–826	1,500–1,200	1,100–850	1,000	1,673
Wendat–Wyandot	835–459			740	

Northern Iroquoian (3965–1283 BC), Cherokee shares a cognate for 'smoking pipe' with Northern Iroquoian languages. Interestingly, however, Seneca and Cayuga do not share a cognate with other Lake Iroquoian languages as might be expected given their recent origin relative to the appearance of smoking pipes in prehistory. All Northern Iroquoian languages included in our study, but not Cherokee, share a cognate for 'corn' (Tables 7 and 8). This pattern of shared cognates suggests that Proto-Iroquoian split into Proto–Northern and Proto–Southern Iroquoian before the introduction of corn, a finding consistent with the earliest reported (Accelerator Mass Spectrometry

Table 6. Cognates for 'bow', 'bow string', and 'smoking pipe' in Iroquoian languages. Cognates were adapted from Mithun (1984). Shaded words are not cognates.

	Stick (bow?)	Bow (~AD 500)	Bow string	Smoking pipe (<1000 BC)
	*-ʔt-	*-(a)ʔęn-	*-rę(h)s-	*-nǫnawę-
Language				
Cherokee		ka:hl²tsaʔ²ti		ka²nʌ̨:²no:²wa
Tuscarora	áʔneh (bow, gun)	unáčreh	urę̀:θeh	uʔnę̀:weh
Wendat	aʔtaʔ	ˌaʔę^{n}da	ˌarę́hsaʔ	anǫdawę
Wyandot		aʔę:^{n}daʔ	yarę́saʔ	yanǫdáwęʔ
Mohawk		aʔʌ́:naʔ	orʌ́hsaʔ	kanʌ́:nawʌ
Oneida		aʔʌ:náʔ	olʌ́:saʔ	kanu:náwʌʔ
Onondaga		aʔę́:naʔ (pole)	kaęsó:tà:ʔ	kanunawę́ʔtaʔ
Cayuga	atóta:ʔ	aʔę́:naʔ (pole)	kaęsota:ʔ	otsokwahtaʔ
Seneca	káeoʔtaʔ (gun)	waʔę:nǫʔ	kanǫ́hsota:aʔ	ashókwahtaʔ
Susquehannock				kanǫ:na

Note: Superscript numbers in examples indicate tone.

Table 7. Comparison of cognates for the word 'corn' associated with the root *-nę̨h-. Shaded words are not cognates shared with Northern Iroquoian languages.

Language	Root	Example	Gloss	Source
Cherokee	-kth-	u^{2}kta	grain, seed	Feeling 1975
Cherokee	selu	se^{23}lu	corn	Feeling 1975
Tuscarora	-nę̨h-	unę́heh	corn	Rudes 1999
Nottoway	-hnę̨h-	ohnehahk	corn	Rudes 1981
Wendat	-nę̨h-	oñnenha	Blé d'inde [corn]	Potier 1748
Wyandot	-nę̨h-	dŭnnę́·ha'	the (grain of) corn	Barbeau 1960
Mohawk	-hnʌ̨h-	tekahnv́hake	two seeds	Michelson 1973
Oneida	-nʌ̨h-	ka·nʌ́heʔ	seed, pit (of a fruit), grain, oats	Michelson and Doxtator 2002
Onondaga	-nę̨h-	onę́haʔ	corn	Woodbury 2003
Cayuga	-nę̨h-	onę́hę'	corn	Mithun and Henry 1982
Seneca	-nę̨-	ʔonɛ́ɔʔ	corn	Chafe 1967: #1155

Note: Superscript numbers in examples indicate tone. Orthography from the original source is presented for the examples of cognates, resulting in some variability in phonetic notation (e.g., glottal stop represented by ʔ, ', or ').

Table 8. Comparison of cognates for the word 'corn' associated with the root *-nę̨hst-. Shaded words are not cognates shared with Northern-Iroquoian languages.

Language	Root	Example	Gloss	Source
Cherokee	-kth-	u^{2}kta	grain, seed	Feeling 1975
Cherokee	selu	se^{23}lu	corn	Feeling 1975
Tuscarora	-nę̨hsn-	unę́hsneh	a grain of wheat, corn, etc., a seed, pit or stone of fruit	Rudes 1999
Susquehannock	-nɛ̨hst-	onæsta	Söd eller Magijz [seed or maize]	Campanius 1696
Wendat	-nɛ̨hst-	oñnensta	graines	Potier 1748
Wyandot	-nɛ̨hst-	nɛ́'sta'	seed	Barbeau 1960
Mohawk	-nʌ̨st-	ó:nvste	corn	Michelson 1973
Oneida	-nʌ̨st-	o·nʌ́steʔ	corn	Michelson and Doxtator 2002

Note: Superscript numbers in examples indicate tone.
The Onondaga cognate, -nɛ̨st-, means 'testicles' (Woodbury 2003).
The similar Cayuga root -nɛ̨hɛ̨hst- means 'whole corn' (Froman et al. 2002).
A similar Seneca root -nɛ̨ɛ̨hsɔ̨hkw- shows up in 'popcorn' (Chafe 1967).
Orthography from the original source is presented for the examples of cognates, resulting in some variability in phonetic notation (e.g., glottal stop represented by ʔ, ', or ').

Table 9. Words for 'bean' in Iroquoian languages. Shaded words are not cognates shared with Five Nations languages.

Language	Root	Example	Gloss	Source
Cherokee	thuya	tu^{23}ya	Beans	Feeling 1975
Tuscarora†	θaheʔ	saugh-he, θáheʔ	peas, bean	Lawson 1709; Rudes 1999
Wendat	-yareʔs-	o$_{ɛ}$ aresa	fève (bean)	Potier 1748
Wyandot	-yareʔs-	dŭyá·re'sa'	the bean	Barbeau 1960
Mohawk	-saheʔt-	osahè:ta	bean	Michelson 1973
Oneida	-saheʔt-	osahé:taʔ	bean	Michelson and Doxtator 2002
Onondaga	-hsaheʔt-	ohsahéʔtaʔ	bean	Woodbury 2003
Cayuga	-hsaheʔt-	osaheʔtaʔ	beans	Froman et al. 2002
Seneca	-sáeʔt-	osáeʔtaʔ	beans	Chafe 1967
Laurentian	sahe	sahe	bean	Mithun 1984

Note: Superscript numbers in examples indicate tone. Orthography from the original source is presented for the examples of cognates, resulting in some variability in phonetic notation (e.g., glottal stop represented by ʔ, ', or ').

†The word for bean may be borrowed from a Five Nations or Laurentian language.

[AMS] dated) macrobotanical evidence of corn east of the Mississippi River (Holding site [American Bottom, Illinois], 2077±70 BP, cal 2σ 350 BC–AD 80) (Crawford et al. 1997, but see Simon 2017). Interestingly, not all Iroquoian languages share a cognate for 'bean' (see Mithun 1984). As might be expected, Cherokee does not share a cognate for 'bean' with Northern Iroquoian, and within Northern Iroquoian the Huronian languages of Wendat and Wyandot do not share a cognate for 'bean' with Tuscarora or the Five Nations languages (Table 9). This pattern of shared cognates suggests that Proto-Iroquoian split into Proto-Northern and -Southern Iroquoian before the introduction of beans to eastern North America around 700 years ago (Hart and Scarry 1999). If we set aside for the moment the fact that Tuscarora shares a cognate for 'bean' with the Five Nation languages not shared with the Huronian languages, this pattern would also suggest that Proto–Lake Iroquoian split into Proto-Huronian and Proto–Five Nations Iroquoian before the introduction of beans. Brown (2010:285–86), after personal communication with Mithun, suggests that the shared cognate for 'bean' among Northern-Iroquoian-speaking groups, including the Tuscarora, is a product of diffusion (borrowing) rather than shared ancestry. This assertion is based primarily on the fact that the divergence date for Tuscarora (2,100 ya, this study, and 2,200 ya, the glottochronological date in Brown 2010) predates by more than a thousand years the earliest evidence of cultivated beans in the eastern United States (ca. 700 ya). This would require diffusion of the word to Tuscarora—and presumably other Coast Iroquoian languages—after adoption of the bean by the Five Nations groups around AD 1300, but before the Tuscarora word for 'bean' was recorded by Lawson (1709:228) in the late 1600s while the Tuscarora were still in North Carolina.

Our analysis of language diversity based on lexical and geographic distances indicates that the area now occupied by the Onondaga-speaking people of the Finger Lakes region of west-central New York is the likely homeland of the Iroquoian language family (Figure 2). Additional analyses of the Northern Iroquoian, Lake Iroquoian, and Five Nations subgroupings have yielded the same result. We also obtained the same results using geographic distances and the lexicostatistical distances based on the percent of shared cognates presented in Hoffman (1959). A closer look at geographic distance and lexical dissimilarity revealed the potential role of geography in the historical development of relationships among Iroquoian languages. We found a strong positive relationship between geographic and lexical distances, with geographic distance explaining about 80% of the variation in lexical distances among languages (Figure 4a). A similar relationship was observed when we compared average lexical and geographic distances (Figure 4b). The results of our Mantel test also indicated a strong and significant relationship between geographic and lexical distance matrices (r_s=0.882, p=0.049). This relationship was not as strong when the analysis was confined to Huronian and Five Nations languages (r_s=0.796, p=0.018), and it was nonsignificant when confined to the Five Nations languages (r_s=0.447, p=0.067).

DISCUSSION

Our results suggest that Proto–Northern and –Southern Iroquoian diverged in west-central New York during the Late Archaic period, around 4,624 ± 1,341 ya (3965–1283 BC), and that Proto–Northern Iroquoian diverged into Proto–Coast Iroquoian and Proto–Lake Iroquoian during the Middle Woodland period around 2,100 ± 609 ya (709 BC–AD 509). The divergence of the Huronian languages also occurred during the Middle Woodland, at around 1730 ± 502 ya (232 BC–AD 772). The two major subgroups of the Five Nations language grouping began to diverge from each other during the Late Woodland, or near the end of what Smith (1997) terms the Transitional Woodland (ca. AD 500 to 900), around 1,163 ± 337 ya (AD 500–1174). Emergence of individual Five Nations and Huronian languages also occurred during the Late Woodland period, between the late twelfth and late fourteenth centuries. Interestingly, our estimated dates are in close accord with glottochronological estimates, particularly those presented by Brown (2010).

We estimated the geographic homeland of the Iroquoian language family using the idea applied by Sapir (1916) that the homeland is typically found in the area of greatest diversity. The results of our analysis of linguistic diversity suggest that all of the major Iroquoian language divergences, with the exception of the split between Wendat and Wyandot, and the splits among the languages of the Coast Iroquoian grouping (i.e., Tuscarora-Nottoway-Meherrin), occurred in the general geographic area historically occupied by the Onondaga people. In other words, the area in and around the Finger Lakes region of west-central New York is the homeland of all the major subgroupings of the Iroquoian language family. Our results support Buell's (1979) determination based on an extensive analysis of Iroquoian cognates for natural history

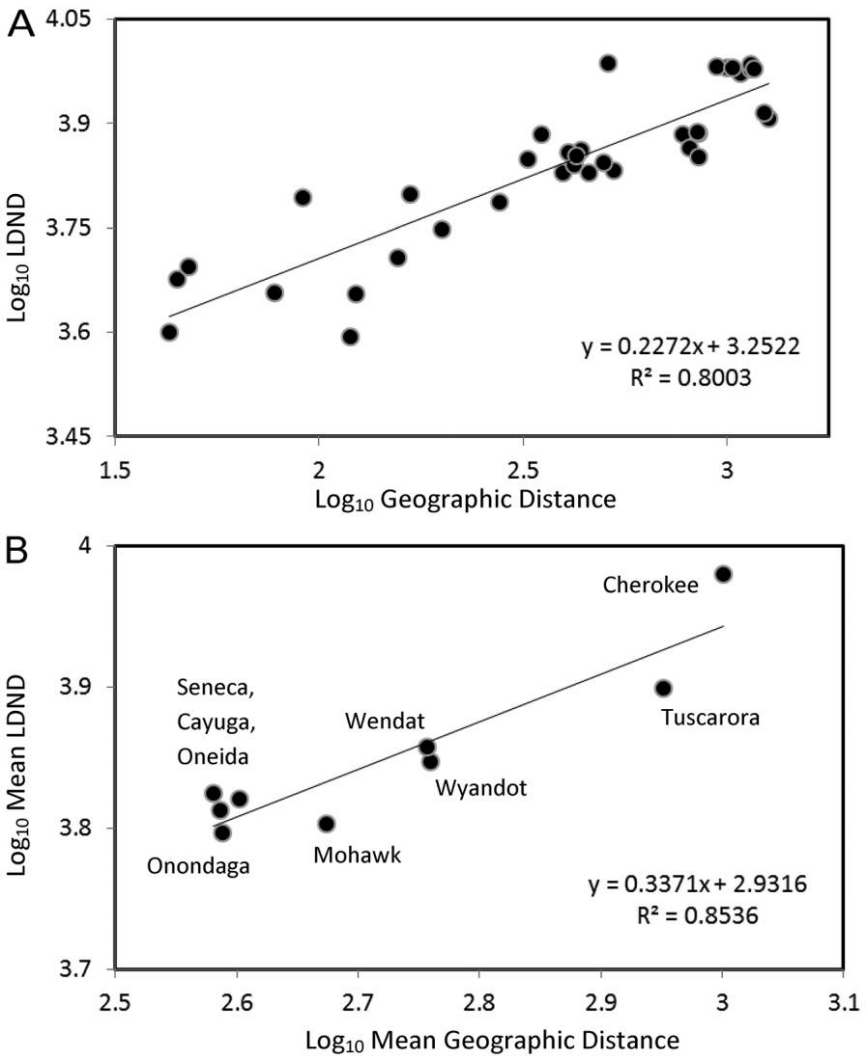

Figure 4. Least-squares linear regression of (A) $\log_{10}$-transformed lexical (LDND) and geographic distance measures among languages and (B) the mean $\log_{10}$-transformed lexical (LDND) and geographic distance measures for each language.

terms for plant and animal species, and the geographic ranges of these species. Her results indicated that the Proto-Iroquoian homeland encompassed the lower Great Lakes region, including southern and eastern Ontario; northern, western, and west-central New York; northwestern Pennsylvania; and northeastern Ohio (Buell 1979:46, map 14). However, our evaluation of the Iroquoian terms for eight tree species presented by

Buell (1979), with the addition of terms presented by others (Christjohn and Hinton 1996; Michelson and Doxtator 2002; Mithun 1984; http://www.cherokeedictionary.net), found that none of the Cherokee terms for these tree species, most of which are found within portions of the geographic area historically occupied by the Cherokee, were cognate with corresponding terms from Northern Iroquoian languages, with the possible exception of the term for *Tsuga canadiensis* (eastern hemlock) (Table 10).

Also, not all Northern Iroquoian languages shared cognate terms for all eight species. Only *Acer saccharum* (sugar maple) and *Ulmus rubra* (slippery elm) exhibited a universally shared cognate term among all Northern Iroquoian languages. *Ulmus americana* (American elm), *Fagus grandifolia* (American beech), and evergreen species such as *Pinus strobus* (white pine), *Pinus resinosa* (red pine), and *Tsuga canadensis* (eastern hemlock) exhibit cognate terms across multiple cognate sets shared variously among Northern Iroquoian languages. There is one universally shared term among Lake Iroquoian languages for *Tilia americana* (American basswood). Although this pattern of shared cognates does not inform us on the location of the Proto-Iroquoian homeland, it does support the suggestion by Buell (1979:46, map 14) that the geographic area she described in her thesis is the Proto–Northern Iroquoian homeland (Figure 5). This suggestion is further supported if the cognate term for the northern species *Larix laricina* (tamarack), which is universally shared among Northern Iroquoian languages, was not developed independently or borrowed (see Mithun 1984:271).

Based on her evaluation of cognates, Mithun (1984:274, 279) indicated that "One semantic complex appears to be quite old. The set of words relating to water suggests that the Proto-Iroquoians lived near a large river or lake" (1984:274) and "The picture that emerges of the Proto-Iroquoians through their vocabulary is one of a people living near water and relying heavily on fishing" (1984:279). Mithun's description seems to echo Buell's results pointing to the lower Great Lakes region. Our estimate for the location of the Iroquoian homeland, as well as that by Buell (1979) and Mithun (1984), is consistent with Lounsbury's (1978:336) discussion of the original homeland of the Iroquoian language family. In this discussion he explains that comparative study of languages shows that those that have undergone the most substantial phonetic changes—which he indicates are first Cherokee and then Tuscarora—from the ancestral protolanguage—said to most closely resemble the Northern Iroquoian languages—are the languages that have moved away from the homeland. He then states, "It would be difficult to interpret the linguistic evidence as favoring a hypothesis of a southern or more westerly origin of the Iroquoian linguistic family. . . . More probably it should be seen as favoring a long occupation of the area of central New York state and north-central Pennsylvania, extending back in time for perhaps as much as four millennia, with expansions or migrations first to the south and then to the north and immediate west" (Lounsbury 1978:336). It is important to note, as did Buell (1979:44), that Proto-Iroquoian speakers did not necessarily inhabit this entire region, but rather could have occupied any portion of it.

Table 10. Iroquoian terms for tree species used to estimate the geographic homeland of Proto-Iroquoian and Proto–Northern Iroquoian. Cognate sets are highlighted in different shades of gray. The various orthographies of the original sources are used, resulting in some variability in phonetic notation (e.g., glottal stop represented by ʔ, ', or '). Sources (provided in footnotes) are indicated by superscripted numbers inside parentheses.

A. Wild plum (*Prunus americana*)

Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
kwanunsdi'i[(1)]		tonestes[(1)]	ta'tun·nę'st[(1)]	oná'onhste'[(1)]	wíhsųʔ[(2)]			kę:eh
kwą²nun³sdį²ʔi		atonnenst	ta'ᵃtuinɛsti[(1)]	(oná'uhste')[(1)]				
			(aʔ)tu:nę́hst					

Note: Wyandot and Wendat terms are cognates, but none of the others.

B. Sugar maple (*Acer saccharum*)

Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
tsu·wa'ki[(1)]	wáht(w)	ouhatta[(1)]	wa'ta'[(1)]	wáhta'[(1)]	wahta'[(1)]	owáhtah[(1)]	wáhta'[(1)]	wahta'[(1)]
tsuwagi		8acta	wáhtaʔ	wáhta	wáhtaʔ	awohátqua[(1)]	ohwáhtaʔ	wahtaʔ
tsʌwaki						ôhwatta[(1)]		
tsûnwagi unega adsilû´skĭ[(3)]						ohwáhtaʔ		

Note: The Lake Iroquoian languages share cognates, which most likely match Tuscarora for 'white ash'; Cherokee does not share a cognate term.

C. Eastern cottonwood (*Populus deltoides*)

Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
a'takvnihi·tv[1]	kahskę?nara?θ?ú?y					guenhnens[1]		
adạ gyṇịhidv								

D. Slippery elm (*Ulmus rubra*)

Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
tsu?ginv[4]	hú'ks[1]	8hŏch	húcuyá:ra'[1]	ohó:ksera[1]	ohò:kseli[1]	onhoóskah[1]	ohóhskra'[1]	oskaia'[1]
dawatsilv[4]	hú?ks		uhšuya:ra?		ohó·ksli?	ohóskæ:?[5]	ohóhskra?	ó:skæ?
dâwātsíla[3]								áoskæ?

Note: Cherokee does not share a cognate term with Northern Iroquoian languages.

E. American beech (*Fagus grandifolia*)

Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
Kusá[6]	utckúrho[1]	ondean		otskén:rha[1]	o?ne·yál	jozgœra[1]	otsɛ'ɛ[1]	oskɛ'ɛ[1]
	učkyérha?	8dêat / 8ndeat				oskenhra[1]	onohotskę?ę?	oskę́?ę?
						otskę́?æ?		

Note: Mohawk, Onondaga, Cayuga, and Seneca share one cognate set, while Oneida and Wendat share another. An old Onondaga term looks cognate with Tuscarora. Cherokee does not share a cognate term with Northern Iroquoian languages.

Table 10 (Continued)

F. American basswood (*Tilia americana*)								
Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
I•tEhá[6]	uhústa[1]		ohochra[1]	ohó:sera[1]	ohósela[1]	hohósa[1]	oho:tra'[1]	o:osæ'[1]
	uhústa'tsæh[1]		hušrá?		ohósle?[7]	hossera[1]	ohó:tra?[5]	o:osæ?[5]
	uhúhsteh				ohó:sela?[2]	ohóhsæ:? [5]		
	uhústra? [5]							
G. American elm (*Ulmus americana*)								
Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
tsu?ginv	karátkwar[1]	˛arak8at	karakõat[1]	oká:ratsi[1]	kalanikwala[1]	kahő:kahta'[1]	khaõka:'[1]	kaõkæ:'[1]
dawatsilv			(u)yaráhkwa?t	a?ká:ratsi		kahú:ka:?	kahọ́:ka:?	kaọkæ:?

Note: Mohawk, Wendat, Wyandot, and Tuscarora share one cognate, Onondaga, Cayuga, and Seneca another. Oneida and Cherokee are different. Cherokee and Oneida do not share cognate terms with other Iroquoian languages.

H. Eastern white pine (*Pinus strobus*), Red pine (*Pinus resinosa*), or Eastern hemlock (*Tsuga canadensis*)								
Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
noh²ji	uhtæhnæh(1)	xahⁿdéhta'(1)	andeta(1)	ohnéhta'(1)	onehta(1)	ôhnetta(1)	onô·'ta'(1)	o:nɛ̃'ta'(1)
notsi(4)	(utéhneh)(1)	andeta	adehta	onʌ́:ta	ohnehta'(1)	ohnéhtaʔ	onę́ʔtaʔ	onę́ʔtaʔ
nona	uhtéhneh	onnenta		onʌ̨ʔtaʔų́:weh	hneht(1)	ohehtah(1)	kanę́ʔtę̨ʔs	oʔsóæʔ
natsi(4)	unʌ̨́ʔteh	oset8ton8a			ohnéhtaʔ		sa:nętáhthaʔ	
to-tsi tlu-gv-i(4)					onʌ̨·taʔ		ostaʔa:ʔ	
atsŭ´nki unega(3)								
tâtskĭ´(3)								

Note: One cognate set (generic pine, loblolly pine) shows up in all the Northern Iroquoian languages except Cayuga and Seneca. The other cognate set (pine, hemlock, fir, evergreen) holds for all the Northern Iroquoian languages except Onondaga, and Wyandot. One of the Cherokee terms may be cognate with the latter.

I. Tamarack (*Larix laricina*)								
Cherokee	Tuscarora	Wendat	Wyandot	Mohawk	Oneida	Onondaga	Cayuga	Seneca
	kanvʔtvs(5)	kannentens(5)		kanv:tvʔs(5)	kanvʔtv:saʔ (5)	kanéʔtɛ̃ʔs(5)	kanéʔtɛ̃ʔs(5)	kanéʔtɛ̃ʔs(5)

Note: Mithun (1984: 271) indicates that the term for 'tamarack' is analyzable as 'falling evergreen', which may reflect the fact that this tree loses its needles. As such these cognate terms for Tamarack, which is found well north of the geographic area historically occupied by the Cherokee, could have developed independently by each group, or alternatively, this term was easily borrowed.

(1) Buell 1979
(2) Michelson and Doxtator 2002
(3) Cozzo 2004
(4) http://www.cherokeedictionary.net
(5) Mithun 1984
(6) Banks 1953
(7) Christjohn and Hinton 1996

The "Algonquian Sea"

Iroquoian-speaking peoples have been described as an "island" in a "sea" of Algonquian-speaking peoples (Williamson 2012:275). As such, the origin, historical development, and geographic distribution of Iroquoian-speaking groups in the southern Great Lakes region all must have been influenced by the historical development of regional Algonquian peoples. Echoing Fiedel's (1991:25) discussion regarding the expansion and diversification of Algonquian languages, any historical scenario explaining the geographic distribution of Iroquoian-speaking peoples at the time of European contact (i.e., the "island") must also explain the geographic distribution of Algonquian peoples.

Fiedel (1987, 1990, 1991) developed a historical model for the development of Algonquian languages based on the results of his glottochronological analysis of Algonquian languages. As a starting point, Fiedel (1987, 1990, 1991) cites Siebert's (1967) determination based on an analysis of Algonquian cognates for plant and animal species names, and the geographic ranges of those species, that the Proto-Algonquian homeland can be found between Lake Ontario and Georgian Bay in southern Ontario (also see Denny 1991).[5] The glottochronological results of Fiedel (1991) indicate that the primary divergence of the Central (e.g., Cree, Ojibwa, Shawnee) and Eastern (e.g., Delaware, Abenaki, Narragansett) Algonquian language groupings occurred around AD 570, though earlier divergence and dispersal of the highly differentiated Micmac (Eastern Algonquian) and Blackfoot (Plains Algonquian) languages is likely. Fiedel (1991) notes that his estimate for the primary divergence of Central and Eastern Algonquian languages corresponds with Lounsbury's (1961) estimate for Proto–Five Nations Iroquois (AD 500–800). He suggests that this correspondence along with a southern Ontario homeland for Proto-Algonquian indicates that an Iroquoian intrusion may have precipitated the breakup of Algonquian languages. In other words, the Iroquoian "island" is the result of an intrusion by Proto–Northern Iroquoian speakers into the southern Great Lakes region inhabited by Proto-Algonquian speakers.

Our chronological estimates for the divergence of Iroquoian protolanguages, and the location of the Proto-Iroquoian homeland, are not necessarily inconsistent with the scenario proposed by Fiedel (1991) for the historical development of Algonquian languages. Based on our determination, the Proto-Iroquoian homeland is located in the Finger Lakes region of west-central New York, well south of Siebert's (1967) estimate for the Proto-Algonquian homeland north of Lake Ontario. After considering potential error, we estimate that Proto–Northern Iroquoian was spoken in the west-central New York homeland during the Early Woodland, between 2624 and 100 BC. Based on Fiedel's glottochronological estimates, Proto-Algonquian would have been spoken within its southern Ontario homeland at about the same time. Also, our estimate for the divergence of Ontario's Huronian languages (i.e., 232 BC–AD 772) encompasses a 1,000-year period during the Middle Woodland which may either coincide with, or precede, Fiedel's estimate for the divergence of Central and Eastern Algonquian languages. Given these estimated protolanguage homelands and divergence dates, it is possible that both language families developed within their homelands

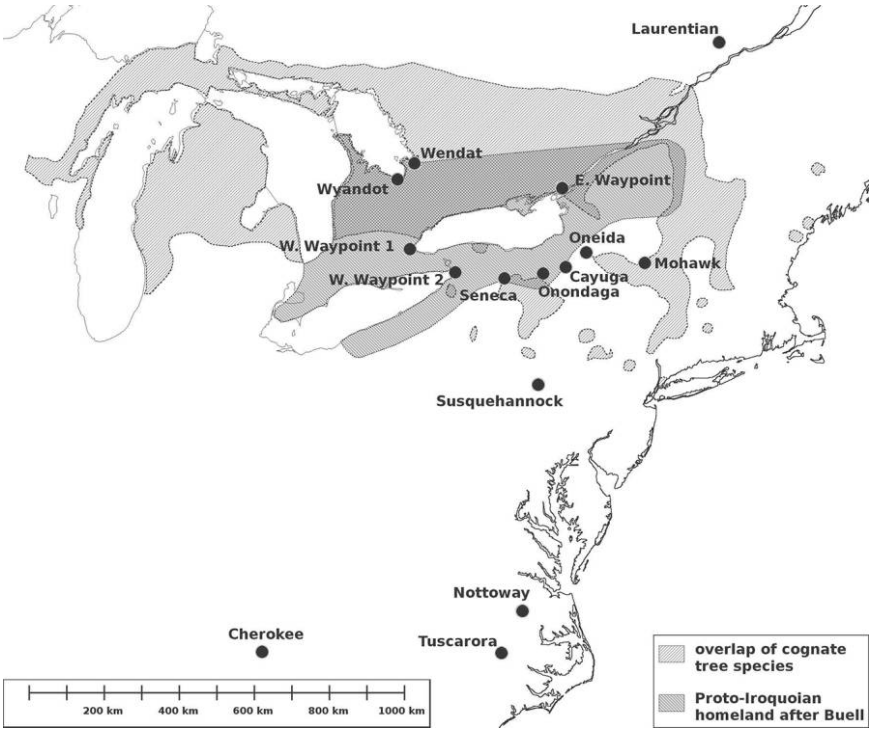

Figure 5. Map showing overlap of the geographic ranges of the tree species with cognate terms (see Table 10). This area of overlap represents the estimated Proto–Northern Iroquoian homeland in the present study. The location of the Proto-Iroquoian homeland presented by Buell (1979:46) is also depicted. The overlap of these estimated homelands is depicted in darker gray.

before expanding. In the case of the Algonquian languages this would include early dispersal of languages such as Blackfoot and Micmac, with subsequent expansions west and east associated with the divergence of Central and Eastern Algonquian languages. For the Iroquoian languages, the divergence of the Huronian languages between 232 BC and AD 772, after the divergence of Cherokee and the Coast Iroquoian languages, would have been precipitated by linguistic isolation from other Northern Iroquoian languages coinciding with the expansion into Ontario—given the geographic location of Huronian-speaking peoples at the time of European contact.

Our finding that the lexical relationships among Iroquoian languages are strongly correlated with geographic proximity is consistent with the predictions of the generalized isolation-by-distance model. This finding, combined with the results of our homeland analysis, suggests that the current geographic distribution of Iroquoian languages, as well as the historical relationships among those languages, developed from a series of dispersals from west-central New York beginning approximately 4,600 years

ago during the Late Archaic period. The first two of these dispersals were to the south (Proto-Cherokee, then Proto–Coast Iroquoian), whereas the third was to the north (Proto-Huronian). Subsequent language divergences among Five Nations languages may simply have been the result of aggregation into regionally distinct populations, perhaps associated with increased sedentism tied to intensification of maize agriculture, rather than dispersal per se. Linguistic innovations would have been shared in a way that was proportional to geographic proximity, with separated speech communities drifting apart until dialects became languages. Although intensive agriculture and population increases would not have been responsible for the initial dispersing protolanguages (i.e., Proto–Coast Iroquoian and Proto-Huronian), their divergence from Proto–Northern Iroquoian would have been attributable to the same isolation-by-distance processes.

Correspondences with the Archaeological Record

As pointed out by Birch (2015), the term "Iroquoian" refers to both language and culture. The essential elements of precontact and historical Iroquoian culture were a primary reliance on maize agriculture, habitation in villages comprising bark-covered longhouses, fortification of villages with palisades, political organization based on village councils, a social organization based on matrilineal descent with clan membership extending beyond the village, nations of affiliated villages, regional confederacies, and ritualized warfare which included trophy taking and prisoner sacrifice (Birch 2015: 265; Williamson 2014:3). It is commonly thought that the full expression of the Iroquoian cultural pattern is not apparent archaeologically until around the turn of the fourteenth century, though Williamson (2014:3, 9) cautions that the emergence of early Iroquoian culture should be viewed as a multilinear process, with differential adoption of settlement and subsistence strategies, as well as social, political, and economic developments, occurring at slightly different times in different regions.

We offer several examples of how our results regarding the timing of language divergences within the Iroquoian language family may correspond with important events observable in the archaeological record within the estimated homeland of the Northern and Southern Iroquoian language groupings. As indicated by Birch (2015), the development and eventual reliance on corn agriculture in the northeastern United States has been considered an Iroquoian trait by archaeologists. Although our estimate for the divergence of Southern and Northern Iroquoian (4,624 ± 1,341 ya, 3965–1283 BC) predates the earliest evidence of corn in eastern North America, our estimate for the divergence of Tuscarora (Coast Iroquoian) from other Northern Iroquoian languages (2,100 ± 609 ya, 709 BC–AD 509) largely coincides with the first phytolith evidence of corn within the estimated Proto-Iroquoian homeland (New York, Vinette site, 2,270 ± 35 BP, cal 2σ 399–208 BC) (Hart et al. 2007) but is older than the earliest macrobotanical evidence (southern Ontario, Grand Banks, 1,570 ± 90 BP, cal 2σ AD 260–660) (Crawford et al. 1997). Our estimate for the divergence of the Huronian languages (Wendat and Wyandot) from Five Nations languages (1730 ±

502 ya, 232 BC–AD 772), however, does overlap with this earliest macrobotanical evidence of corn from the Grand Banks site in southern Ontario. It is important to note here that our analysis of language diversity based on lexical and geographic distances indicated that all of these primary divergences would have occurred within west-central New York. This suggests that Proto–Northern Iroquoian groups in New York would have knowledge of and, importantly, share a word for corn, whereas Proto–Southern Iroquoian groups (i.e., Cherokee) would not.

Our analysis of presumed cognates for 'corn' seems to confirm this expectation. As pointed out by Mithun (1984: 272), the corn complex in Iroquoian languages presents an interesting puzzle, and as such, identifying cognates for 'corn' is not straightforward. There is an overlap in two Proto–Northern Iroquoian roots which both have reflexes meaning 'seed' or 'corn,' where the meaning varies across the languages. For instance, the root *-nęh- means 'seed' in Mohawk and Oneida but 'corn' in the other languages, while *-nęhst- means 'corn' in Mohawk and Oneida but 'seed' in the others (Mithun 1984:272). Compare the cognate sets of each morpheme presented in Table 7. The column labeled "Root" indicates the morpheme in question—in other words, the smallest component of the word bearing that meaning. The full-word example and its gloss, or meaning, are taken directly from each indicated source and thus show variant orthographies. The reflexes of *-nęh- are straightforwardly cognate in form, while the different meanings (seed vs. corn) split Mohawk and Oneida from the rest of the family. The reflexes, or descendant forms, of *-nęhst- in Table 8 also separate Mohawk and Oneida from the other languages, with a reversal of the meaning difference. Unlike in Table 7, here Onondaga, Cayuga, and Seneca do not match up precisely with the others. Neither of these two Northern Iroquoian forms is cognate with either the Cherokee word for seed or for corn.

Anthropologists have long considered longhouses—multifamily residential structures that have a width-to-length ratio of 1 : >2 (Hart 2000; Kapches 1984)—to be an important Iroquoian trait. Cross-cultural research has shown that residential structures that exceed 79.2 m^2, as most Iroquoian longhouses do, can be used as a threshold to infer matrilocal postmarital residence with 95% confidence (Divale 1977; also see Ember 1973; Brown 1987). More recently, research by Porčić (2010) has generally confirmed the earlier results, with the caveat that a significant relationship between floor area and matrilocal residence only holds for agricultural societies. Matrilocality also seems to be associated with recent migration or external warfare (Divale 1977) and may even have played a role in the evolution of corn stocks in the study area (see Hart 2001). Research compiling prehistoric house sizes and associated radiocarbon dates (AMS) has indicated that the development of longhouses may have begun as early as the mid AD 1000s (based on 2σ ranges) and became more frequent in the 1200s (Hart 2000), with those exhibiting floor areas indicative of matrilocal residence not common until the 1100s (Hart 2001). Longhouses seem to develop earlier in Ontario, perhaps as early as the mid-900s, though the organization of interior space may suggest only incipient matrilocality at that time (Kapches 1990).

Our results suggest that during the development of longhouses, and presumably matrilocal postmarital residence, in New York, Proto-Onondaga/Cayuga/Seneca split into Proto-Onondaga and Proto-Seneca/Cayuga languages (822 ±238 ya, AD 940–1460), and soon thereafter Proto-Seneca/Cayuga split into Proto-Seneca and Proto-Cayuga (795 ±232 ya, AD 973–1437).

Our dates for the split of Proto-Oneida/Mohawk into Proto-Mohawk and Proto-Oneida fall within the range estimated for the development of corn-bean-squash cultivation in New York. A recent review of radiocarbon (AMS) dates associated with the occurrence of beans at Iroquoian sites in New York suggest that corn-bean-squash cultivation, an indicator of agricultural intensification (Crawford 2014), may have begun as early as the mid-to-late 1200s (based on 2σ ranges) and became established in the 1300s in both New York and Ontario (Hart 2001; Hart and Scarry 1999).

Limitations

The results of our study are subject to several important assumptions and limitations. Our method of estimating language split dates assumes the approximate correctness of the 52 calibration dates used by ASJP (Holman et al. 2011) in support of the assumption of roughly equal rates of lexical change across languages. Lexical similarities among the Iroquoian languages are assumed to be attributable to common linguistic ancestry, rather than innovations shared or borrowed during (re)contact. Chafe and Foster (1981) have asserted that the well-attested relationship between Seneca and Cayuga is in fact partly a product of diffusion between Seneca and Cayuga during multiple periods of recontact after initial separation (also see discussion in Brown 2010 and Julian 2010). Their careful analysis of shared innovations groups Cayuga with Tuscarora, rather than other Five Nations languages. If their analysis is correct, our reconstruction of the historical relationships among Iroquoian languages, and the timing of their divergences, is at least partly incorrect. Similarly, as we have pointed out, the word for 'bean' within Northern Iroquoian may also be a product of diffusion. Our study is also potentially limited by the ASJP's standardized orthography that utilizes a limited set of phonological symbols and does not take into account, for example, vowel length or tone. Phonological attributes such as vowel length and tone, however, are often not included in the historical documentation of a language, and how such attributes are heard and recorded are more likely to vary among linguists. It is important to note, however, that our tree describing the relationships among Iroquoian languages is identical to those based on two separately determined cognate lists (i.e., Hoffman 1959; Julian 2010) and the pattern of mutual intelligibility among languages, as well as very similar to trees presented by Iroquoian experts (cf. Lounsbury 1978; Mithun 1984).

Our study is also limited by the availability of reasonably accurate and complete lexical data. Such data were not available for a number of Iroquoian languages known historically, including Susquehannock, Wenro, Neutral, Erie, Nottoway, Meherrin,

and the Laurentian languages. In addition, other languages may have gone extinct prior to European contact, or were never described. A reviewer of this paper pointed out that our study cannot accommodate the extinction of possible languages that once existed geographically between Cherokee, the outlier within the Iroquoian language family, and surviving Northern Iroquoian languages. This limitation, of course, applies to all studies of the historical relationships among languages.

CONCLUSION

Our results suggest that the initial breakup of Proto-Iroquoian into Proto–Southern and Proto–Northern Iroquoian dates to 4,624 ± 1,341 ya (3965–1283 BC), and that the Finger Lakes region of west-central New York is the most likely Proto-Iroquoian homeland. This homeland is the geographic region inhabited by Late Archaic Proto-Iroquoian speakers immediately prior to the divergence of Proto-Northern and -Southern Iroquoian sometime between 3965 and 1283 BC. The geographic origin and the areal extent of the region inhabited by Proto-Iroquoian speakers in the centuries before the initial language divergence, as well as their regional and extraregional population movements (e.g., migration from the south, or west), however, is unknown. The results of the Mantel test and regression analyses revealed a strong relationship between linguistic dissimilarity and geographic distance, which, based on a generalized isolation-by-distance model, we interpret as reflecting the isolating effects of spatial separation on the magnitude of linguistic exchange. In other words, the relationships among Iroquoian languages reflect a historical process of dispersals and relative isolation which has led to lects drifting apart, eventually becoming separate languages. The timing of language divergences seems to coincide sometimes with important events observable in the archaeological record. The emergence of the Northern Iroquoian grouping is roughly associated with the first evidence of the use of corn in New York, and the first evidence of corn in Ontario roughly coincides with the emergence of Proto-Huronian. This observation is supported by the pattern of shared cognates for 'corn'. The development of important Iroquoian cultural attributes such as the longhouse, matrilocal residence, and the intensification of agriculture all become established in the archaeological record at about the same time that most of the internal language divergences within the Lake Iroquoian language grouping (i.e., Five Nations and Huronian languages) occurred. Based on these coincidences we speculate that the structure of language relationships within the family may be attributable to the consequences of maize agriculture, including population growth, sedentism, and population aggregation (see Kohler et al. 2008) leading to relative isolation and lects drifting apart.

The major contribution of the present research is its multi-method approach to the study of Iroquoian language prehistory, one that included a comprehensive statistical analysis of lexical and geographic data, along with an evaluation of select cognates. We also highlight the potential of linguistic data for providing clues to the prehistory of Native American/First Nations groups. The statistical analysis of lexical

data offers archaeologists an independent and complementary tool for investigating regional cultural histories.

NOTES

The title of this article was inspired by Irvine Davis's (1959) paper on Kiowa-Tanoan prehistory. We are grateful to K. Michelson, C. Dyck, J. Steckley, P. Abrams, and A. Vick, who contributed lexical data or advice. We thank J. Bischoffberger (Leiden University) for transcribing and entering the lexical data into the ASJP program, and H.-J. Bibiko and H. Pinto for their help with maps. This paper benefited greatly from discussions with R. Williamson, G. Crawford, and D. Smith. S. Wichmann's research was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant Agreement No. 295918, and by a subsidy of the Russian government to support the Program of Competitive Development of Kazan Federal University.

1. We use the Early, Middle, and Late Woodland periods as largely arbitrary, though commonly used, temporal units. We do not mean to imply these temporal units represent bounded cultural units or archaeological cultural taxa (see discussion in Hart 2011).

2. http://www.ciera.ulaval.ca/English/publications/yawenda-en.htm

3. Orthographies were regularized by using a consistent set of symbols, particularly in addressing representation of affricates and allophonic voicing of certain consonants. Because the status of Wendat and Wyandot as different languages or merely dialects of one language is in question, each was kept separate using data only from clearly attested sources. Since Wendat records were phonemically deficient (e.g., glottal stop not recorded), the Wendat data were partially reconstructed to cover the deficient areas. This may have artificially rendered the Wendat data even closer to Wyandot than it actually was.

4. Software made available by J. Relethford (http://employees.oneonta.edu/relethjh/programs/).

5. The languages of the Algonquian subfamily, along with the Wiyot and Yurok languages found in present-day California, comprise the Algic language family. Although the homeland of the Algonquian subfamily, or portions of the subfamily, may be found in southern Ontario or elsewhere in northeastern United States, there is disagreement regarding the location of the homeland of the larger Algic language family, which may instead be found to the west, either on the Columbia Plateau or along the northwestern coast of North America (see discussions in Goddard 1994; Sapir 1916; Snow 1976; Wichmann et al. 2010).

REFERENCES CITED

Banks, William H. 1953. *Ethnobotany of the Cherokee Indians*. Master's thesis, University of Tennessee.

Barbeau, C. Marius. 1960. *Huron-Wyandot traditional narratives in translations and native texts*. Anthropological Series 47, National Museum of Canada Bulletin 165. Ottawa.

Bateman, Richard, Ives Goddard, Richard O'Grady, V. A. Funk, Rich Mooi, W. John Kress, and Peter Cannell. 1990. Speaking of forked tongues: The feasibility of reconciling human phylogeny and the history of language. *Current Anthropology* 31:1–24.

Birch, Jennifer. 2015. Current research on the historical development of Northern Iroquoian societies. *Journal of Archaeological Research* 23:263–323.

Blin-Lagarde, Pierrette. 1972. *Une étude historique dans les langues de la famille huronne-iroquoise*. Master's thesis, McGill University, Montreal.

Blitz, John H. 1988. Adoption of the bow in prehistoric North America. *North American Archaeologist* 9:123–45.

Boyce, Douglas W. 1978. "Iroquoian tribes of the Virginia–North Carolina coastal plain," in *Handbook of North American Indians 15: Northeast*. Edited by Bruce G. Trigger, pp. 290–95. Washington, DC: Smithsonian Institution.

Brown, Barton M. 1987. Population estimation from floor area: A restudy of "Naroll's Constant." *Cross-Cultural Research* 21:1–49.

Brown, Cecil. H. 2006. "Glottochronology and the chronology of maize in the Americas," in *Histories of maize: Multidisciplinary approaches to the prehistory, biogeography, domestication, and evolutions of maize*. Edited by John E. Staller, Robert H. Tykot and Bruce F. Benz, pp. 647–63. San Diego: Elsevier.

———. 2010. "Prehistoric chronology of the common bean in the New World: The linguistic evidence," in *Pre-Columbian foodways: Interdisciplinary approaches to food, culture, and markets in ancient Mesoamerica*. Edited by John E. Staller and Michael Carrasco, pp. 273–91. New York: Springer Science.

Buell, Nancy S. 1979. *The Proto-Iroquoian homeland*. Master's thesis, State University of New York at Albany.

Campanius, Johannes. 1696. *Lutheri catechismus öfwersatt på American-Virginiske språket*. Stockholm: J.J. Ganath.

Campbell, Lyle. 2015. Do languages and genes correlate? Some methodological issues. *Language Dynamics and Change* 5:202–26.

Chafe, Wallace L. 1967. *Seneca morphology and dictionary*. Smithsonian Contributions to Anthropology 4. Washington, DC: Smithsonian Press.

Chafe, Wallace L., and Michael K. Foster. 1981. Prehistoric divergences and recontacts between Cayuga, Seneca, and the other Northern Iroquoian languages. *International Journal of American Linguistics* 47:121–42.

Christjohn, Amos, and Maria Hinton. 1996. *An Oneida dictionary / Ukwehu·wehneha Tekawʌnate?nyése*, edited by Cliff Abbot. Oneida, WI: Oneida Tribe of Indians of Wisconsin.

Chrisomalis, Stephen, and Bruce G. Trigger. 2004. "Reconstructing prehistoric ethnicity: Problems and possibilities," in *A passion for the past: Papers in honour of James F. Pendergast*. Edited by J. V. Wright and Jean-Luc Pilon, pp. 419–34. Mercury Series Archaeology Paper 164. Gatineau: Canadian Museum of Civilization.

Cozzo, David N. 2004. *Ethnobotanical classification system and medical ethnobotany of the Eastern Band of the Cherokee Indians*. PhD dissertation, University of Georgia, Athens.

Crawford, Gary W. 2014. "Food production and niche construction in pre-contact southern Ontario," in *Reassessing the timing, rate, and adoption trajectories of domesticate use in the*

Midwest and Great Lakes. Edited by Maria E. Raviele and William A. Lovis, pp. 135–60. Occasional Papers 1. Midwest Archaeological Conference. Available online at http://www.midwestarchaeology.org/storage/MAC-Occasional-Papers-No-1.pdf

Crawford, Gary W., and David G. Smith. 1996. Migration in prehistory: Princess Point and the Northern Iroquoian case. *American Antiquity* 61:782–90.

Crawford, Gary W., David G. Smith, and Vandy E. Bowyer. 1997. Dating the entry of corn (*Zea mays*) into the lower Great Lakes. *American Antiquity* 62:112–19.

Davis, Irvine. 1959. Linguistic clues to northern Rio Grande prehistory. *El Palacio* 66:73–84.

Denny, J. Peter. 1991. "The Algonquian migration from Plateau to Midwest: Linguistics and archaeology," in *Papers of the 22nd Algonquianist Conference*. Edited by William Cowan, pp. 103–34. Ottawa: Carleton College.

Divale, William T. 1977. Living floor area and marital residence: A replication. *Cross-Cultural Research* 12:109–15.

Ember, Melvin. 1973. An archaeological indicator of matrilocal versus patrilocal residence. *American Antiquity* 38:177–82.

Feeling, Durbin. 1975. *Cherokee-English dictionary*. Tahlequah: Cherokee Nation of Oklahoma.

Fiedel, Stuart J. 1987. Algonquian origins: A problem in archaeological-linguistic correlation. *Archaeology of Eastern North America* 15:1–12.

———. 1990. Middle Woodland Algonquian expansion: A refined model. *North American Archaeologist* 11:209–30.

———. 1991. Correlating archaeology and linguistics: The Algonquian case. *Man in the Northeast* 41:9–32.

Froman, Frances, Alfred Keye, Carrie J. Dyck, and Lottie Keye. 2002. *English-Cayuga/Cayuga-English dictionary*. Toronto: University of Toronto Press.

Goddard, Ives. 1994. "The west-to-east cline in Algonquian dialectology," in *Actes du Vingt-Cinquième Congrès des Algonquinistes*. Edited by William Cowan, pp. 187–211. Ottawa: Carleton University.

Gooskens, Charlotte, Wilbert Heeringa, and Karin Beijering. 2008. Phonetic and lexical predictors of intelligibility. *International Journal of Humanities and Arts Computing* 2:63–81.

Hart, John P. 2000. New dates from classic New York sites: Just how old are those longhouses? *Northeast Anthropology* 60:1–22.

———. 2001 Maize, matrilocality, migration, and Northern Iroquoian evolution. *Journal of Archaeological Method and Theory* 8:151–82.

———. 2011. "The death of Owasco—redux," in *Current Research in New York State Archaeology: A.D. 700–1300*. Edited by Christina B. Rieth and John P. Hart, pp. 95–108. New York State Museum Record 2. Albany: State University of New York.

Hart, John P., and William Engelbrecht. 2012. Northern Iroquoian ethnic evolution: A social network analysis. *Journal of Archaeological Method and Theory* 19:322–49.

Hart, John P., and Margaret C. Scarry. 1999. The age of common beans (*Phaseolus vulgaris*) in the northeastern United States. *American Antiquity* 64:653–58.

Hart, John P., Hetty Jo Brumbach, and Robert Lusteck. 2007. Extending the phytolith evidence for early maize (*Zea mays* ssp. *mays*) and squash (*Curcurbita* sp.) in central New York. *American Antiquity* 72:563–83.

Hickerson, Harold, Glen D. Turner, and Nancy P. Hickerson. 1952. Testing procedures for estimating transfer of information among Iroquois dialects and languages. *International Journal of American Linguistics* 18:1–8.

Hoffman, Bernard G. 1959. Iroquois linguistic classification from historical materials. *Ethnohistory* 6:160–85.

Holman, Eric W., Søren Wichmann, Cecil H. Brown, Viveka Velupillai, André Müller, Pamela Brown, and Dik Bakker. 2008. Explorations in automated language classification. *Folia Linguistica* 42:331–54.

Holman, Eric W., Cecil H. Brown, Søren Wichmann, André Müller, Viveka Velupillai, Harald Hammarström, Sebastian Sauppe, Hagen Jung, Dik Bakker, Pamela Brown, Oleg Belyaev, Matthias Urban, Robert Mailhammer, Johann-Mattis List, and Dmitry Egorov. 2011. Automated dating of the world's language families based on lexical similarity. *Current Anthropology* 52:841–75.

Holman, Eric W., Søren Wichmann, Cecil H. Brown, and E. Anthon Eff. 2015. Inheritance and diffusion of language and culture: A comparative perspective. *Social Evolution and History* 14:49–64.

Julian, Charles. 2010. *A history of the Iroquoian languages*. PhD dissertation, University of Manitoba.

Kapches, Mima. 1984. Cabins on Ontario Iroquois sites. *North American Archaeologist* 5:63–71.

———. 1990. The spatial dynamics of Ontario Iroquoian longhouses. *American Antiquity* 55: 49–67.

Kohler, Timothy A., Matt Pier Glaude, Jean-Pierre Bocquet-Appel, and Brian M. Kemp. 2008. The Neolithic demographic transition in the US Southwest. *American Antiquity* 73:645–69.

Konigsberg, Lyle W. 1990. Analysis of prehistoric biological variation under a model of isolation by geographic and temporal distance. *Human Biology* 62:49–70.

Kopris, Craig A. 2001. *A grammar and dictionary of Wyandot*. PhD dissertation, State University of New York at Buffalo.

Lawson, John. 1709. A new voyage to Carolina; containing the exact description and natural history of that country, together with the present state thereof and a journal of a thousand miles travelled through several nations of Indians, giving a particular account of their customs, manners, etc. London [no publisher].

Lees, Robert B. 1953. The basis of glottochronology. *Language* 29:113–27.

Lounsbury, Floyd G. 1961. "Iroquois-Cherokee linguistic relations," in *Symposium on Cherokee and Iroquois culture* (Vol. 180), edited by William N. Fenton and John Gulick, pp. 11–17. Bureau of American Ethnology Bulletin. Washington DC: Smithsonian Institution.

———. 1978. "Iroquoian languages," in *Handbook of North American Indians, 15: Northeast*. Edited by Bruce G. Trigger, pp. 334–43. Washington, DC: Smithsonian Institution.

Mantel, Nathan. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209–20.

Martin, Scott W. J. 2008. Languages past and present: Archaeological approaches to the appearance of Northern Iroquoian speakers in the Lower Great Lakes region of North America. *American Antiquity* 73:441–63.

Michelson, Günther. 1973. *A thousand words of Mohawk*. Ethnology Division Paper 5. Ottawa: National Museums of Canada.

Michelson, Karin, and Mercy Doxtator. 2002. *Oneida-English/English Oneida dictionary*. Toronto: University of Toronto Press.

Mithun, Marianne. 1981. Stalking the Susquehannocks. *International Journal of American Linguistics* 47:1–26.

———. 1984. "The proto-Iroquoians: Cultural reconstruction from lexical materials," in *Extending the rafters: Interdisciplinary approaches to Iroquoian studies*. Edited by Michael K. Foster, Jack Campisi, and Marianne Mithun, pp. 259–81. Albany: State University of New York Press.

Mithun, Marianne, and Reginald Henry. 1982. *Watewayęstanih: A Cayuga teaching grammar*. Brantford: Woodland Indian Cultural Educational Centre.

Ortman, Scott G. 2012. *Winds from the North: Tewa origins and historical anthropology*. Salt Lake City: University of Utah Press.

Porčić, Marko. 2010. House floor area as a correlate of marital residence pattern: A logistic regression approach. *Cross-Cultural Research* 44(4):405–24.

Potier, Pierre. 1748. Vocabulaire huron-français. Ms. collection Gagnon. Montreal Public Library.

Rafferty, Sean M. 2016. "Smoking pipes of Eastern North America," in *Perspectives on the archaeology of pipes, tobacco and other smoke plants in the ancient Americas*. Edited by Elizabeth A. Bollwerk and Shannon Tushingham, pp. 13–26. Cham, Switzerland: Springer.

Rudes, Blair A. 1981. A sketch of the Nottoway language from a historical-comparative perspective. *International Journal of American Linguistics* 47: 27–49.

———. 1999. *Tuscarora-English/English-Tuscarora dictionary*. Toronto: University of Toronto Press.

Sapir, Edward. 1916. *Time perspective in aboriginal American culture: A study in method*. Anthropological Series No. 90-92. Ottawa: Government Printing Bureau.

Saitou, Naruya, and Masatoshi Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406–25.

Schillaci, Michael A., and Wendy J. Bustard. 2010. Controversy and conflict: NAGPRA and the role of biological anthropology in determining cultural affiliation. *Political and Legal Anthropology Review* 33:352–73.

Siebert, Frank T. 1967. "The original home of the Proto-Algonquian people," in *Contributions to Anthropology: Linguistics I (Algonquian)*. Edited by A. Donald DeBlois, pp. 13–47. Ottawa: National Museum of Canada, Bulletin No. 214, Anthropological Series No. 78.

Simon, Mary L. 2017. 2017. Reevaluating the evidence for Middle Woodland maize from the Holding site. *American Antiquity* 82:140–50.

Smith, David G. 1997. Radiocarbon dating the Middle to Late Woodland transition and earliest maize in southern Ontario. *Northeast Anthropology* 54:37–73.

Snarey, Kristen, and Christopher Ellis. 2008. "Evidence for bow and arrow use in the Smallpoint Late Archaic of southwestern Ontario," in *Papers in honor of Michael Spence*. Edited by Neal Ferris, Peter Timmins, Christine D. White, and Chris J. Ellis. Occasional Publication of the London Chapter, No. 9. *Ontario Archaeology* 85–88:21–38.

Snow, Dean R. 1976. "The archaeological implications of the proto-Algonquian *urheimat*," in *Papers of the Seventh Algonquianist Conference, 1975*. Edited by William Cowan, pp. 339–46. Ottawa: Carleton University.

———. 1995. Migration in prehistory: The Northern Iroquoian case. *American Antiquity* 60: 59–79.

———. 1996. More on migration in prehistory: Accommodating new evidence in the Northern Iroquoian case. *American Antiquity* 61:791–96.

———. 2009. "The multidisciplinary study of human migration," in *Ancient human migrations: A multidisciplinary approach*. Edited by Peter N. Peregrine, Ilia Peiros, and Marcus Feldman, pp. 6–20. Salt Lake City: University of Utah Press.

Stothers, David M. 1977. *The Princess Point complex*. Archaeological Survey of Canada, Mercury Series Paper 58, Ottawa: National Museum of Man.

Swadesh, Morris. 1950 Salish internal relationships. *International Journal of American Linguistics* 16:157–67.

———. 1955. Towards greater accuracy in lexicostatistical dating. *International Journal of American Linguistics* 21:121–37.

Tamura, Koichiro, Glen Stecher, Daniel Peterson, Alan Filipsky, and Sudhir Kumar. 2013. MEGA 6: Molecular evolutionary genetics analysis, version 6.0. *Molecular Biology and Evolution* 30:2725–29.

Tang, Chaoju, and Vincent J. van Heuven. 2009. Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119:709–32.

Trigger, Bruce G. 1970. The strategy of Iroquoian prehistory. *Ontario Archaeology* 14:3–48.

———. 1978. "Introduction," in *Handbook of North American Indians 15: Northeast*. Edited by Bruce G. Trigger, pp. 1–3. Washington, DC: Smithsonian Institution.

Warrick, Gary. 2000. The precontact Iroquoian occupation of southern Ontario. *Journal of World Prehistory* 14:415–66.

Wichmann, Søren, André Müller, and Viveka Velupillai. 2010. Homelands of the world's language families: A quantitative approach. *Diachronica* 27:247–76.

Wichmann, Søren, Eric W. Holman, and Cecil H. Brown, eds. 2016. The ASJP database (version 17). http://asjp.clld.org/.

Williamson, Ronald F. 2012. "What will be has always been: The past and present of Northern Iroquoians," in *The Oxford handbook of North American Indians*. Edited by T. R. Pauketat, pp. 273–84. Oxford: Oxford University Press.

———. 2014. The archaeological history of the Wendat to A.D. 1651: An overview. *Ontario Archaeology* 94:3–64.

Woodbury, Hanni. 2003. *Onondaga-English/English-Onondaga dictionary*. Toronto: University of Toronto Press.

Wright, James. 1984. "The cultural continuity of the Northern Iroquoian-speaking peoples," in *Extending the rafters: Interdisciplinary approaches to Iroquoian studies*. Edited by Michael K. Foster, Jack Campisi, and Marianne Mithun, pp. 283–99. Albany: State University of New York Press.

Wright, Sewall. 1943. Isolation by distance. *Genetics* 28:114–38.